Docket No.: 08228/1203278-US1
(PATENT)

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Letters Patent of:
Sharon Flank et al.

Patent No.: 7,177,879

Issued: February 13, 2007

For: METHOD AND APPARATUS FOR DIGITAL MEDIA MANAGEMENT, RETRIEVAL, AND COLLABORATION

REQUEST FOR CERTIFICATE OF CORRECTION
PURSUANT TO 37 CFR 1.323 AND 1.322

Attention: Certificate of Correction Branch
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

Upon reviewing the above-identified patent, Patentee noted typographical errors which should be corrected. A listing of the errors to be corrected is attached.

The typographical errors marked with an "A" on the attached list are found in the application as filed by applicant. Please charge our Credit Card in the amount of $100.00 covering the fee set forth in 37 CFR 1.20(a).

The typographical errors marked with a "P" on the attached list are not in the application as filed by applicant. Also given on the attached list are the documents from the file history of the subject patent where the correct data can be found.

The errors now sought to be corrected are inadvertent typographical errors the correction of which does not involve new matter or require reexamination.

Transmitted herewith is a proposed Certificate of Correction effecting such corrections. Patentee respectfully solicits the granting of the requested Certificate of Correction.

The Commissioner is authorized to charge any deficiency of up to $300.00 or credit any excess in this fee to Deposit Account No. 04-0100.

Dated: March 13, 2007

Respectfully submitted,

By _____

Flynn Barrison
Registration No.: 53,970
DARBY & DARBY P.C.
P.O. Box 5257
New York, New York 10150-5257
(212) 527-7700
(212) 527-7701 (Fax)
Attorneys/Agents For Applicant

PTO/SB/44 (04-05)
Approved for use through 04/30/2007. OMB 0651-0033
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

(Also Form PTO-1050)

UNITED STATES PATENT AND TRADEMARK OFFICE
CERTIFICATE OF CORRECTION

PATENT NO. : 7,177,879 Page 1 of 1

APPLICATION NO.: 10/063,410

ISSUE DATE : Feb. 13, 2007

INVENTOR(S) : Flank et al.

It is certified that an error appears or errors appear in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

On the face page, in field (75), in "Inventors", in column 1, line 2, delete "Washington, DC" and insert - - Reston, VA - -, therefor.

In column 8, line 52, delete "addition," and insert - - In addition, - -, therefor.

In column 9, line 32, after "If it is" delete "riot," and insert - - not, - -, therefor.

In column 12, line 28, after "system then" delete "then".

MAILING ADDRESS OF SENDER (Please do not use customer number below):
Flynn Barrison
DARBY & DARBY P.C. 1
P.O. Box 5257
New York, New York 10150-5257

Issued Patent Proofing Form File#: **08228/1203278-US1**
Note: **P** = PTO Error **A** = Applicant Error

US Serial No.: **10/063,410** US Patent No.: **US 7,177,879 B2** Issue Dt.: **Feb. 13, 2007**
Title: **METHOD AND APPARATUS FOR DIGITAL MEDIA MANAGEMENT, RETRIEVAL, AND COLLABORATION**

Sr. No.	P/A	Original		Issued Patent		Description of Error
		Page	Line	Column	Line	
1	P	Page 3 Oath or Declaration filed (07/22/2002)	1 (Second Inventor's Residence)	First Page Col. 1 (Inventors)	2	Delete "Washington, DC" and insert - - Reston, VA - -, therefor.
2	A	Page 13 of 34 Specification (04/21/2002)	16	8	52	Delete "addition," and insert - - In addition, - -, therefor.
3	P	Page 14 of 34 Specification (04/21/2002)	25	9	32 (Approx.)	After "If it is" delete "riot," and insert - - not, - -, therefor.
4	A	Page 19 of 34 Specification (04/21/2002)	22	12	28 (Approx.)	After "system then" delete "then". (Second Occurrence)

US007177879B2

(12) **United States Patent**
Flank et al.

(10) **Patent No.: US 7,177,879 B2**
(45) **Date of Patent: Feb. 13, 2007**

(54) **METHOD AND APPARATUS FOR DIGITAL MEDIA MANAGEMENT, RETRIEVAL, AND COLLABORATION**

(75) Inventors: **Sharon Flank**, Vienna, VA (US); **Keith Simonsen**, Washington, DC (US); **Robert Griffin**, Vienna, VA (US)

(73) Assignee: **Corbis Digital Asset Management Services, LLC**, Seattle, WA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 782 days.

(21) Appl. No.: **10/063,410**

(22) Filed: **Apr. 21, 2002**

(65) **Prior Publication Data**

US 2002/0194198 A1 Dec. 19, 2002

Related U.S. Application Data

(60) Division of application No. 10/063,409, filed on Apr. 21, 2002, now Pat. No. 6,922,691, which is a continuation of application No. PCT/US01/26841, filed on Aug. 28, 2001.

(60) Provisional application No. 60/228,837, filed on Aug. 28, 2000.

(51) **Int. Cl.**
G06F 7/00 (2006.01)
G06F 17/00 (2006.01)
G06K 9/00 (2006.01)

(52) **U.S. Cl.** **707/104.1**; 382/126; 715/526

(58) **Field of Classification Search** 382/126; 715/526
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

5,532,839	A *	7/1996	Beikirch et al.	358/401
5,600,775	A	2/1997	King et al.	
5,682,487	A *	10/1997	Thomson	715/800
5,721,902	A	2/1998	Schultz	
5,752,244	A	5/1998	Rose et al.	
5,850,561	A	12/1998	Church et al.	
5,864,845	A	1/1999	Voorhees et al.	
5,938,724	A	8/1999	Pommier et al.	
5,983,237	A	11/1999	Jain et al.	
5,987,459	A *	11/1999	Swanson et al.	707/6
6,006,241	A	12/1999	Purnaveja et al.	
6,038,333	A *	3/2000	Wang	382/118
6,072,904	A	6/2000	Desai et al.	

(Continued)

FOREIGN PATENT DOCUMENTS

WO	WO 02/19147	3/2002

Primary Examiner—Jeffrey Gaffin
Assistant Examiner—Cindy Nguyen
(74) *Attorney, Agent, or Firm*—Darby & Darby PC; Jamie L. Wiegand

(57) **ABSTRACT**

The software according to the invention incorporates a glossary management tool that makes it easy for each client to customize terminology to the needs of a particular business. With this tool, termed a glossary manager, a company can customize a number of feature names in the system to provide a more familiar context for their users. A system administrator can also customize the manner in which "thumbnail" or "preview" images are presented. The system performs clustering on search queries, and searches media records multi-modally, using two or more approaches such as image searching and text searching. An administrator can tune search parameters. Two or more streams of metadata may be aligned and correlated with a media file, facilitating later searching. The system evaluates itself. It folds popularity information into rankings of search results.

11 Claims, 4 Drawing Sheets

Questions Associated with RUP 129

Question Id	Question	Distributes Over	Role
106	new question2	Groups	Modifier

and if two or more parts of speech are possible for a particular word, it is tagged with both. After tagging, word affixes (i.e. suffixes) are stripped from query words to obtain a word root, using conventional inflectional morphology. If a word in a query is not known, affixes are stripped from the word one by one until a known word is found.

An intermediate query is then formulated to match against the file index database. Texts or captions that match queries are then returned, ranked, and displayed to the user, with those that match best being displayed at the top of the list. In an exemplary system, the searching is implemented by first building a B-tree of ID lists, one for each concept in the text database. The ID lists have an entry for each object whose text contains a reference to a given concept. An entry consists of an object ID and a weight. The object ID provides a unique identifier and is a positive integer assigned when the object is indexed. The weight reflects the relevance of the concept to the object's text, and is also a positive integer.

To add an object to an existing index, the object ID and a weight are inserted into the ID list of every concept that is in any way relevant to the text. For searching, the ID lists of every concept in the query are retrieved and combined as specified by the query. Since ID lists contain IDs with weights in sorted order, determining existence and relevance of a match is simultaneous and fast, using only a small number of processor instructions for each concept-object pair.

Search Technologies. The system allows users to search for media files with many different types of search queries. For example, users may submit search queries by speaking them, typing them, copying them, or drawing them.

The process of locating a particular file in a large archive is a special area for innovation within the inventive software. Files are characterized in several ways. First, they have an identifier, generally similar to a filename, which is unique within the system and makes it possible to link up all the objects related to a file. These can include the actual high-resolution asset, lower-resolution thumbnails or other proxies for browsing, and information about the file, or metadata. Searching can be performed on the file identifier, or it can be performed on the metadata. In the case of metadata searching, it is desirable to offer search alternatives that go beyond the exact matching process involved in a standard keyword search.

Some systems use controlled vocabulary searching as an optimization of keyword searching. Keyword searches simply match exactly on any word in the user's search query that appears in the search target. (In the system according to the invention, the search target is the metadata describing a media file.) The set of potential keywords is quite large (as large as the vocabulary of English, or whatever language(s) are being used). If there are no limitations on the search vocabulary that can be employed, a user can enter a search for puma and fail to find any files captioned as mountain lion or cougar, even though they all refer to the same thing. Controlled vocabulary is an attempt to address this problem, albeit at considerable cost. In a controlled vocabulary retrieval system, cataloguers all agree to use the same terms. In practical terms, this implies that, when cataloguing, they must check their controlled vocabulary lists and be sure not to deviate. Sometimes tools can be built to aid in this process, depending on the size of the controlled vocabulary. Similarly, tools can also be provided to searchers to control their search requests. However, controlled vocabulary systems do not scale beyond a few thousand terms, since it is impractical to look up every word in English for every search. For broader retrieval systems, for faster cataloguing, and for simpler searching, a different approach is superior.

In addition to standard keyword and Boolean searching, the system software incorporates additional advanced technology for locating stored files. Rather than limiting searching to a controlled vocabulary, the system software includes natural language search, which allows cataloguers and users to employ any words in English (or whatever natural language the retrieval system is using).

Natural language search incorporates:

a semantic network of concepts

additional linguistic techniques, including:

phrase matching

derivational morphology, in lieu of stemming

part of speech tagging

name recognition

location recognition

User-tunable Search Parameters. The system according to the invention provides a screen for customers to adjust search parameters, to reflect their company use of stored media file collections. This is shown in FIG. **3**. While the parameters may themselves be well-known in a searching system, what is emphasized here is that the user (or, more likely, an administrator) can be granted access to such fundamental decisions about search as:

(a) how good a match has to be before it is displayed to the user, e.g. 50%, and

(b) how "creative" the search should be, i.e. how much should the search terms be expanded to include more distant synonyms and related terms.

It should be borne in mind that the system according to the invention can be carried out on an internet, meaning an IP-based network, and in particular may be carried out on the Internet, meaning the global IP-based network.

Multimodal Search. Currently, search methods focus on textual input. The current invention incorporates new search techniques, and combines them in novel ways.

Image search is becoming useful in commercial applications. In the system according to the invention, user search input is provided in a new way. Users may wish to select an existing image as example input, so that a search consists of "Give me more images like this." Perhaps even more useful is the ability to select part of an image, analogous with "Give me more like this part." In the system according to the invention, identifying the part may be done in either of two exemplary ways:

1. Touch screen: user touches the screen to identify the portion of the image that feeds into the search.

2. Markup, using pen or other screen drawing metaphor, including through the system media viewer, which is described in more detail below.

addition, search modalities can be combined. This novel approach to search is particularly applicable to multimedia. Examples of combined, or multimodal, searches, include:

voice and text

touch screen and text

drawing and voice

drawing and touch screen

and so on.

Vocabulary Management. Ideally, the semantic net of concepts is quite large and attempts to incorporate every word or term in English (or other language being used for cataloguing and searching). No matter how large the semantic net may be, there will be a periodic need to expand or edit it. New words appear in English periodically, and, although many may be slang and therefore not particularly important in a business context, some will be real new words and will

be important enough to include. For example, rollerblading and in-line skating are relatively new terms in English, and depicting those actions is useful in advertising. So the terms need to be added to the semantic net. Semantic net/vocabulary maintenance is generally a manual process, particularly where the user has an existing media library with a thesaurus and vocabulary management process. Such maintenance can also be performed automatically.

To maintain a vocabulary for an information retrieval application that accepts user queries in natural language, a user maintaining a semantic net would track search queries in a query log. From the query log, he would determine which words are actually novel and are candidates to be added to the system vocabulary, by expanding the query log using morphology, and possibly a spell checker and name identifier. The remaining terms that were not matched are the basis of a list for adding terms to the vocabulary.

Tools to manage vocabulary include:

A morphological analyzer. This tool strips off any endings and morphological alterations in a query to find the stem, and checks to see if the stem is in the current vocabulary. If the stem is not, the user doing the maintenance might try:

A spell checker. This tool uses the conventional algorithms to see if the supposedly new word is actually a misspelling of a known word. If it is not a misspelling, the user might try:

A name identifier. This tool checks to see if the supposedly new word is in a name configuration, in that it follows a known first name in the query. If it does, it is added to a candidate name database. If it is riot, it is proposed as a possible new word to be added to the system's vocabulary.

Searching Audio/Video by Timecode Correlation with Search Criteria. Video and audio files can be timecoded, or marked such that the software in which they run can locate a specific frame (for videos) or measure (for audio) at any time. Importantly, the system according to the invention permits searching timemedia, including video and audio files, by combining two search elements. The first is a standard search, including but not limited to natural language search. The second is a time indicator, such as a SMPTE (Society of Motion Picture and Television Engineers) standard timecode. Face recognition is an additional technology that can be used in searching. Face recognition is a subset of the more general technology of object recognition, and indeed techniques described here may extend to additional technologies as well.

The current state of the art in face recognition technology makes it possible to take a manually created, labeled library of faces, and match faces from a video to that library. For example, a user might work with a news video and use a face recognition program to label Nelson Mandela in it. The output of the face recognition program would be a time-coded segment, with start and stop times, of when Nelson Mandela was on camera, with the label "Nelson Mandela" attached to the time codes. While face recognition currently does not achieve 100% precision or recall, it can still be useful. For example, one known system offers a contract rights management capability for films that demands time-coded segments with names attached, and assumes that users will create those manually, so that the correct contract restrictions for each film segment can be attached to the right time codes. Given a small library of the actors in a film, it would be possible to do a fast, automated match-up of time codes and actors, even with imperfect face recognition technology. Selecting the correct actor from forty publicity shots would be much simpler than selecting from among thousands of faces.

Importantly, the system according to the invention carries out the automated creation of the face library. Required elements include time-coded metadata (for example, the voice recognition transcript of a video), and the ability to find the names of people in text. Each time a face and a person's name appear at the same time code, that occurrence is a potential new entry for the face library. A user may run the facematcher for thousands of hours and sift out the recurring matches as the most likely. In this way, a reference library of faces is created, and new material can be catalogued automatically.

The software according to the invention approaches this by using alignment techniques to match up two or more streams of metadata. For example, a broadcast news program may contain closed captioning for the hearing-impaired. It may also contain a separate description of the news footage, probably created manually by the news department. The system according to the invention uses alignment to match the description, which is not time-coded, with the closed captioning, which is time-coded. This process allows the system to add time codes to the non-time-coded stream. The software then uses that new, derived stream (i.e. the description with newly added time codes), and searches for proper names within it. At the same time, using face recognition algorithms on the video stream, the software finds faces. The system tries to match up the faces with the proper names that describe who they are. This matched set provides us with a rough cut of a face (or object) reference library. This is exemplified in FIG. 4.

Face recognition can also be employed to manage the library or archive of media files. Media libraries are assembled over time, often from disparate sources, and may contain multiple copies of a single media file, either with the same metadata or with different metadata. Duplicate detection is therefore an important element of library and archive management, and face recognition (and, more generally, image recognition) can be leveraged to provide that capability. More broadly, for video, scene detection technology can assist in the process of identifying duplicates so that they can be purged from the library.

Clustering and Other Ways to Determine Stored File Usage. Clustering involves combining user search queries in such a way that the searches can be analyzed usefully to provide answers to business questions. Clustering has received considerable attention in document information retrieval (IR) and more recently, in video IR as a means of refining retrieval results based on user preferences or profiles, and to characterize the marketplace. The prior art contains many examples of clustering applied in information retrieval systems, but they all apply to search results returned to users rather than search queries submitted by users. In the system according to the invention, we cluster search queries by topic. We then use that information to adjust the collections of stored files so that the file collections will better meet users' needs.

This system characterizes the information needs of groups (and subgroups) of users, with respect to a collection of media files (e.g. images, videos, sound clips, text, multimedia objects). Some common groupings include:

search queries that brought back no files

search queries that brought back no files the user was interested in

search queries that lead to expressions of user interest or sales

This system applies clustering technology to user-submitted search queries, and to the files retrieved in search results. It also includes:

- machine learning as applied to the above.
- characterizing the information needs over time, by user type, or by other factors.
- methods for reporting file collection needs to interested parties (for example, media suppliers). The system informs a supplier that pictures of earthquakes are selling briskly, or that users keep looking for videos of dance performances but cannot find any.
- methods for adjusting file collections based on the results of the clustering analysis, above.
- novel clustering techniques. These include using a semantic expansion (such as the WordNet hierarchical thesaurus) and phrase identification (such as noun phrase, and name and location identification) as the basis for the clustering.

Before user queries can be analyzed, they must be expanded to a "common denominator". To expand the user search queries, we use natural language techniques. Specifically, we treat each query as if it were metadata within our system, as described in the NLP section, above. Each query is expanded through the application of a semantic net, so that it contains synonyms and related terms. This expansion provides a set of features to which we can apply standard clustering technology.

Obtaining valuable information on user preferences for stored files begins with deciding what information a client wants to understand. The data set can be selected according to various criteria, including:

- queries from a particular subset of users (e.g. registered users, users by industry, new users)
- queries that lead to success (sale or other indication)
- queries that lead to failure

A first step is to select the data set on which clustering is to be performed. In an information retrieval (IR) context, clustering can be performed on queries or on assets to be retrieved (documents, images, video, audio, mixed media). A sample query set may include short queries, as is standard on Web searches, long queries, as seen in TREC (U.S. government-sponsored text retrieval conferences), or as produced by QBE (query by example), in which an asset or set of assets are themselves inserted into a query.

A second step is to perform analysis on the queries using, for example, linguistic methods, such as:

- Tokenization: determine word/token boundaries. In English, tokenization mostly coincides with spaces between words, with certain complications (Alzheimer's=1 token vs. she's=2 tokens).
- Morphology or stemming: removed tense and plural markers and other affixes to find the word root.
- Identify names, locations, noun phrases: using a pattern matcher or other methodology, determine words were groupings for special handling. For example, for names, match certain kind of variance; for locations, match subset; for noun phrases, we to complete and headmatches higher than modifier homematches.

A third step is to expand the queries. Ideally, this step includes expansion using a thesaurus or semantic net of synonyms, superand other relationships.

A fourth step is, for each of the terms in each expanded query, assign a weight based on how close that term is to the original query. The exact weighting will vary by application, but the basic understanding is that more closely matching terms are weighted close to 100.

A fifth step is to create a vector for each expanded query. In order to apply a statistical clustering algorithm, we arrange the vectors into a matrix.

A sixth step is to apply a statistical clustering algorithm in order to group similar queries together. A hierarchical or linear clustering strategy may be used, depending on whether the desired clusters are hierarchical or not. Clustering may allow overlap, which means that a query may appear in more than one cluster.

A seventh step is to apply the clustering algorithm until the stopping condition is met, e.g. the desired number of clusters is obtained, or a combination of cluster number and desired distinctiveness of clusters is reached.

An eighth step relates to the clusters. Clusters are most useful to a human observer if they bear useful names that reflect the content. Use a semantic net hierarchy, combined with term frequency in a reference corpus, to identify the lowestterm in the hierarchy that subsumes all the queries.

Folding Popularity Into Rankings. Many information retrieval applications currently incorporate relevance feedback into their judgements of how to rank search results returned to a user. In all cases, however, the past systems utilize explicit user feedback, not implicit feedback. That is, they rank files by requiring a user to indicate what items he is interested in, once a set of items is returned by a search. Importantly, in the system according to the invention, the system discerns implicit popularity rankings based on a ranked set of user actions. The system then then uses those rankings to resubsequent search results. The user actions from which popularity may be determined include, but are not limited to:

- whether a file is placed in a projects folder or other work space
- whether a file is placed in a shopping cart
- whether a file is purchased

In addition, implicit popularity rankings may be derived from one user or set of users and applied to a different user or set of users. For example, if User A places a media file in a projects folder or shopping cart, the information on her activity can also be used to research results for User B, who is in some way similar to user A. In another example, if users with ".edu" e-mail addresses buy certain things, it makes sense to research results to favor those things when showing results to other ".edu" users. In the system according to the invention, if registered users who work for advertising agencies have placed certain items in their shopping carts, other advertising agency employees can have their search results re-ranked to favor those items. The opposite can be true as well: the same system can be used to disfavor certain items because they have been sold too many times to advertising agencies, for example.

Retrieval system self-evaluation. The retrieval system according to the invention does not answer any of the TREC tracks—a media file is described by a short paragraph or by keywords. Even though it is possible to develop a unique test collection for purposes of evaluation, it will not necessarily predict the performance of new systems, or even for existing ones.

What is now described is a method for ongoing evaluation of IR system performance based on search results combined with user feedback. This approach enables the system to alert a human system manager about observed degradation in the system performance. As the performance evaluation system becomes more knowledgeable through methods of machine learning, the system is desirably able to change its own parameters in order to improve its performance.